昆 虫 学 报 *Acta Entomologica Sinica*, November 2017, 60(11): 1235 - 1246 doi:10.16380/j.kcxb.2017.11.001

意大利蝗转录组及性腺发育相关基因分析

向 敏[1], 叶小芳[1], 扈鸿霞[1], 王伟亮[2], 于 非[1],
肖宏伟[2], 季 荣[1,*], 王 晗[1,*]

(1. 新疆特殊环境物种多样性应用与调控实验室, 中亚区域跨境有害生物联合控制国际研究中心,
新疆师范大学生命科学学院, 乌鲁木齐 830054;
2. 新疆玛纳斯县蝗虫鼠害预测预报防治站, 新疆玛纳斯 832200)

摘要:【目的】意大利蝗 *Calliptamus italicus* 是新疆荒漠、半荒漠草原优势危害种类之一,其生殖相关基因信息缺乏。本研究对意大利蝗精巢和卵巢组织进行高通量转录组测序,旨在揭示意大利蝗转录组特征,并获得与性腺发育相关的基因数据。【方法】利用 Illumina 高通量测序平台(HiSeq/MiSeq)对意大利蝗进行了转录组测序,并进行序列分析。【结果】获得 718 872 条 unigenes,平均长度 631 bp, N50 为 1 186 bp。通过同源性比对,成功注释 254 597 条 unigenes,其中,Nr 数据库注释占 28.87%,比例最高。Nr 数据库注释的 unigenes 与黑褐草蚁 *Lasius niger* 的相似基因序列最多,为 10.80%。与 GO 数据库比对发现,意大利蝗转录组 unigene 可分为 3 大类:涉及分子功能的 unigenes 有 31 392 条,涉及细胞组分的 unigenes 有 20 586 条,与生物学过程有关的 unigenes 有57 014条。KEGG pathways 分析表明,23 666 条 unigenes 归属于 251 个通路。涉及生殖系统发育的通路有卵母细胞减数分裂、胰岛素信号通路、Wnt 信号通路、促性腺激素释放激素信号通路、转化生长因子 β 信号通路以及泛素介导的蛋白水解、黄体酮介导的卵母细胞成熟、昆虫激素生物合成等。进一步筛选得到部分与性腺发育相关的基因,包括 *vg*, *vgr*, *sox 21b*, *testis-specific serine/threonine-protein kinase*, *spermatogenesis-associated protein* 和 *vasa-like gene* 等。【结论】本研究获得了意大利蝗转录组数据及性腺发育相关基因,为进一步研究意大利蝗生殖调控机理奠定了分子基础。

关键词: 意大利蝗; 高通量测序技术; 基因注释; 卵巢; 精巢

中图分类号: Q966 **文献标识码:** A **文章编号:** 0454-6296(2017)11-1235-12

Analysis of the transcriptome and gonadal development related genes of *Calliptamus italicus* (Orthopera: Acrididae)

XIANG Min[1], YE Xiao-Fang[1], HU Hong-Xia[1], WANG Wei-Liang[2], YU Fei[1], XIAO Hong-Wei[2], JI Rong[1,*], WANG Han[1,*] (1. Key Laboratory of Special Environment Biodiversity Application and Regulation in Xinjiang, International Research Center for the Collaborative Containment of Cross-Border Pests in Central Asia, College of Life Sciences, Xinjiang Normal University, Urumqi 830054, China; 2. Manasi Station of Forecast and Control on Locusts and Rodents, Manasi, Xinjiang 832200, China)

Abstract:【Aim】*Calliptamus italicus* is a major plague pest in the arid and semi-arid grasslands of Xinjiang, with a lack of information of its reproduction related genes. The objective of this study is to reveal the characteristics of the transcriptome of *C. italicus* and to obtain the gonada development related

基金项目: 新疆维吾尔自治区重点实验室开放课题(2015KL028); 新疆特殊环境物种多样性应用与调控实验室资助项目(XJTSWZ-2017-01); 国家自然科学基金项目(31260104); 科技部国际合作专项(2015DFR30290); 新疆维吾尔自治区天山创新团队计划项目
作者简介: 向敏, 女, 1992 年 8 月生, 新疆库尔勒人, 硕士研究生, 研究方向为昆虫生殖生态学, E-mail: 1441109005@qq.com
* 通讯作者 Corresponding authors, E-mail: wanghanguoxi@sina.com; jirong@xjnu.edu.cn
收稿日期 Received: 2017-07-24; 接受日期 Accepted: 2017-09-13

genes in *C. italicus* through the transcriptome sequencing of its testicular and ovarian tissues. 【Methods】 The transcriptome of *C. italicus* was sequenced using the Illumina HiSeq/MiSeq platform and bioinformatically analyzed. 【Results】 A total of 718 872 unigenes were obtained, with a mean length of 631 bp and an N50 of 1 186 bp. Through a similarity search, 254 597 unigenes were annotated. Most unigenes (28.87%) were annotated to Nr database, and the unigenes of *C. italicus* had the highest similarity (10.80%) with those of *Lasius niger* annotated in Nr database. According to GO database, all unigenes were broadly annotated into three categories: 31 392 unigenes to be related to molecular function, 20 586 unigenes related to cellular component, and 57 014 unigenes related to biological processes. In KEGG database, a total of 23 666 unigenes were assigned to 251 metabolic pathways, of which many pathways are involved in the reproductive system development, including oocyte meiosis, insulin signaling pathway, Wnt signaling pathway, GnRH signaling pathway, TGF-β signaling pathway, ubiquitin mediated proteolysis, progesterone-mediated oocyte maturation, and insect hormone biosynthesis. By further screening and identification, gonadal development related genes of *C. italicus*, including genes of *vg*, *vgr*, *sox 21b*, *testis-specific serine/threonine-protein kinase*, *spermatogenesis-associated protein* and *vasa-like gene*, were obtained. 【Conclusion】 This study acquired the transcriptome data and gonadal development related genes of *C. italicus*, providing a molecular foundation for further studying the mechanism of reproductive regulation of *C. italicus*.

Key words: *Calliptamus italicus*; high-throughput sequencing technology; gene annotation; ovary; testis

随着新一代高通量测序技术的发展,转录组分析已经成为非模式昆虫发掘功能基因的有效手段之一,广泛应用于遗传进化、害虫防治、生态适应等研究中(Oppenheim *et al*., 2015; Zhang *et al*., 2016)。近些年,高通量测序技术被大量应用于蝗虫转录组学研究中(Kim *et al*., 2016; Qiu *et al*., 2016, 2017)。金永玲等(2015)对大垫尖翅蝗 *Epacromius coerulipes* 进行了转录组分析,共得到了 63 033 条 unigenes,筛选得到与杀虫剂抗药性相关的功能基因及潜在 SSR 位点。张硕(2015)通过构建亚洲小车蝗 *Oedaleus asiaticus* 雌雄成虫触角转录组文库,鉴定出编码气味分子结合蛋白基因,并对其表达谱进行测定分析,为进一步研究亚洲小车蝗嗅觉感受的分子机制提供数据基础。

蝗虫具有较强的生殖适应能力,种类多、数量大、分布广,是危害农牧业生产安全的世界性重大害虫,其生殖调控机制一直是国内外研究热点(赵卓和奚耕思, 2005)。蝗虫始终保持较高的产生配子的能力,是其暴发成灾的重要原因(赵卓等, 2008)。性腺是蝗虫繁殖的重要场所,精巢和卵巢发育成熟并行使正常的生理功能,才能产生有效精子和卵子,保障蝗虫种群繁衍(赵卓等, 2008)。因此,对蝗虫性腺发育的分子调控机制开展研究,有利于深入了解蝗虫的生殖调控机理,为害虫的综合防治提供理论支持及数据基础。转录组技术为从分子水平研究昆虫性腺发育和配子发生的调控机制提供了新思路。杨婧(2013)对短额负蝗 *Actractomorpha sinensis* 雌雄成虫转录组进行分析,筛选出 1 879 个雌雄差异基因,发现组织蛋白酶在雌虫体内表达上调,而 CYP3A 在雄虫体内表达上调,推测二者分别参与了卵子发生过程和性激素生物合成过程,为深入研究蝗虫性别分化及生殖调控机理提供了数据支持。

意大利蝗 *Calliptamus italicus* 是新疆荒漠半荒漠草原的优势危害种类,随着全球气候变化,意大利蝗灾害在新疆频繁发生,给新疆的畜牧业经济、生态和社会造成严重的损失(薛智平等, 2010)。目前,对意大利蝗的研究主要集中在生态分布、生长发育、行为特征、生理生化、分子生物学及防治等方面,而关于意大利蝗生殖调控方面的研究相对较少(王晗等, 2010; 赵忠伟等, 2013; Hunter *et al*., 2016; Popova *et al*., 2016; Ren *et al*., 2016; 王冬梅等, 2016; 张晓红等, 2016)。前期研究发现,高温对意大利蝗雌虫繁殖力有显著影响,随温度升高,意大利蝗卵巢发育和卵子发生进程加快(向敏等, 2017)。然而,其分子机制还有待于进一步研究。意大利蝗为非模式生物,可参考的相关基因、蛋白序列较少,因此,本研究借助高通量测序技术对意大利蝗精巢和卵巢转录组进行分析,以期获得卵巢发育相关的重要功能基因,为从分子水平上阐明意大利蝗的生殖适应机制及暴发成灾规律提供科学依据。

1　材料与方法

1.1　供试蝗虫

意大利蝗于 2016 年 7 月采自昌吉回族自治州玛纳斯县(43°54′N, 86°21′E,海拔 1 310 m),带回实验室后置于人工气候箱(浙江托普 GTOP-260B)内,在 27 ± 1℃、相对湿度 45% ± 5%、光周期 14L∶10D 条件下饲养,挑选性成熟期(未交配)、健康的雌雄成虫,分别取卵巢和精巢后迅速投入液氮冷冻,并于 -80℃保存待用。每 3 个卵巢或精巢混合为一个检测样品,卵巢和精巢各共取 3 个检测样品。

1.2　总 RNA 的提取与检测

采用 Trizol Reagent(Invitrogen)提取总 RNA,分别使用 Nanodrop2000(IMPLEN, CA, USA), Qubit RNA Assay Kit, Agilent 2100(Agilent Technologies, CA, USA)和 1.0% 琼脂糖凝胶电泳检测总 RNA 样品的纯度、浓度和完整性。

1.3　cDNA 文库构建和测序

RNA 样品检测合格后,用带有 Oligo(dT)的磁珠富集 mRNA,随后加入 fragmentation buffer 将 mRNA 打断成短片段,以 mRNA 为模板构建测序文库。文库构建完成后,使用 Qubit2.0, Agilent 2100 和 Q-PCR 方法对文库进行检测,以保证文库质量。最后使用 Illumina 高通量测序平台(HiSeq/MiSeq)进行测序。cDNA 文库构建与测序由北京康普森生物技术有限公司协助完成。

1.4　转录组数据的拼接与组装

将测序得到的原始数据(raw reads)经过处理可获得高质量的测序数据(clean data)。同时,计算 Q20 和 Q30, GC 含量和重复序列水平,并利用 Trinity 软件对获得高质量的 clean data 数据进行序列拼接、过滤和组装,最终获得高质量的 unigene。

1.5　基因功能注释

使用 BLAST 软件将 unigene 序列与下列数据库比对,获得基因的功能注释信息。数据库分别为：NCBI 蛋白数据库(NCBI non-redundant protein sequences, Nr), NCBI 核酸序列数据库(NCBI non-redundant nucleotide sequences, Nt), 蛋白家族(protein family, Pfam), 直系同源基因簇(clusters of orthologous groups of proteins, KOG/COG), Swiss-Prot 数据库(a manually annotated and reviewed protein sequence database),京都基因与基因组百科全书(KEGG ortholog database, KO),基因本体论(Gene Ontology, GO)。

1.6　性腺发育相关基因的筛选

根据雌雄样本差异表达基因的功能注释信息,结合已有相关文献报道,筛选并鉴定与性腺发育过程相关的功能基因。

2　结果

2.1　意大利蝗转录组的测序和组装

采用 Illumina 测序平台对意大利蝗转录组进行测序,每个样本生成超过 12.7 Gb 的高质量数据(NCBI 登录号：SRR6113309, SRR6113310)。GC 含量在 41.92% ~47.25% 之间。每个样本 Q20 均大于 97%,Q3 均大于 93%(表 1)。测序结果准确度较高,可以用于后续分析。

利用 Trinity 软件进行序列组装,共得到 832 098 个转录本,总长度为 651 690 832 bp, N50 和平均长度分别为 1 930 bp 和 783 bp。对转录本进行聚类和组装分析,得到 718 872 条 unigenes, N50 和平均长度分别为 1 186 bp 和 631 bp(表 2)。转录组长度分布见表 2,其中长度小于 600 bp 的 unigenes 所占比例最大,为 80.481%。

表 1　意大利蝗 6 个 cDNA 样品测序数据评估统计

Table 1　Summary of sequencing data of six cDNA samples of *Calliptamus italicus*

	样品编号 Sample no.	原始 Reads 数 Number of raw reads	高质量 Reads 数 Number of clean data	数据大小 Size of clean bases (Gb)	GC 含量(%) GC content	Q20 (%)	Q30 (%)
雌虫 Female locust	1	52 459 346	52 270 701	15.68	47.25	97.49	93.58
	2	42 766 547	42 598 957	12.78	41.92	97.28	93.34
	3	43 370 123	43 219 365	12.96	42.20	97.25	93.26
雄虫 Male locust	4	53 694 335	53 500 693	16.05	45.09	97.33	93.32
	5	52 711 023	52 505 418	15.75	44.33	97.27	93.21
	6	55 590 859	55 378 554	16.61	44.37	97.30	93.28

Q20, Q30：分别表示质量值≥20 和 30 的碱基所占百分比 Percentages of bases with the quality value≥20 and 30, respectively.

表 2　意大利蝗转录组组装结果

Table 2　Illumina transcriptome assembly results for *Calliptamus italicus* transcriptome

长度范围 Length range (bp)	转录本 Transcripts 数量 Number	转录本 Transcripts 百分比 Percentage	Unigenes 数量 Number	Unigenes 百分比 Percentage
< =600	610 470	73.365	578 553	80.481
600 – 1 500	161 714	19.434	115 378	16.050
1 500 – 3 000	40 111	4.821	177 66	2.471
3 000 – 6 000	16 351	1.965	6 023	0.838
6 000 – 15 000	3 388	0.407	1 131	0.157
> 15 000	64	0.008	21	0.003
总数 Total number	832 098		718 872	
总长度 Total length (bp)	651 690 832		453 313 494	
N50 长度 N50 length (bp)	1 930		1 186	
平均长度 Mean length (bp)	783		631	

N50：覆盖 50% 所有核苷酸的最大 Unigene 长度 Length of the smallest contig which when added to a set of larger contigs yields at least 50% of the genome.

2.2　基因功能注释

使用 BLAST 软件将获得的 unigenes 序列与 Nr, Nt, Pfam, KOG/COG, Swiss-Prot, KO 和 GO 各数据库进行比对,获得意大利蝗转录组 unigene 的注释信息。在组装获得的 718 872 条 unigenes 中,成功注释 254 597 条 unigenes,占总 unigenes 数的 35.42%。其中,Nr 数据库注释 207 555 条 unigenes,占 28.87%,比例最高,其后依次是 KOG(15.01%), Swiss-Prot (11.10%), Nt(7.69%), GO(5.90%), KEGG (3.29%)和 Pfam(0.03%)(表 3)。

Unigene 注释到 Nr 数据库中的物种分布见图 1,其中与黑褐草蚁 *Lasius niger* 的相似基因序列最多,所占比例为 10.80%,其后依次是内华达古白蚁

表 3　不同数据库中意大利蝗转录组 unigenes 注释结果统计

Table 3　Annonation result statistics of unigenes of *Calliptamus italicus* transcriptome in different databases

注释数据库 Annotated database	被注释的基因数量 Number of annotated unigenes	注释百分比 Percentage of annotated unigenes
Nr	207 555	28.87
KOG	107 886	15.01
Swiss-Prot	79 770	11.10
Nt	55 300	7.69
GO	42 401	5.90
KEGG	23 666	3.29
Pfam	189	0.03
All	254 597	35.42

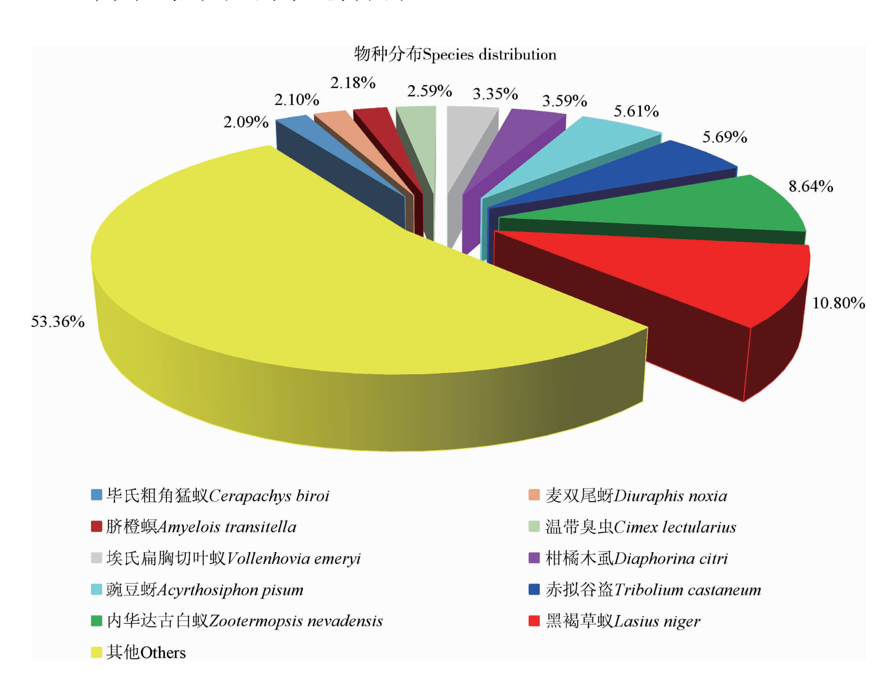


图 1　意大利蝗 unigene 在 Nr 数据库中的物种分布图

Fig. 1　Species distribution of unigenes of *Calliptamus italicus* in Nr database

Zootermopsis nevadensis（8.64%）、赤拟谷盗 *Tribolium castaneum*（5.69%）、豌豆蚜 *Acyrthosiphon pisum*（5.61%）、柑橘木虱 *Diaphorina citri*（3.59%）、埃氏扁胸切叶蚁 *Vollenhovia emeryi*（3.35%），其他相似性大于 2% 的物种有温带臭虫 *Cimex lectularius*（2.59%）、脐橙螟 *Amyelois transitella*（2.18%）、麦双尾蚜 *Diuraphis noxia*（2.10%）、毕氏粗角猛蚁 *Cerapachys biroi*（2.09%），其他物种占 53.36%。

2.3　Unigene 的功能分类

GO 是国际标准化的基因功能分类系统，能够将基因划分为 3 大类，分别是分子功能（molecular function）、细胞组分（cellular component）和生物过程（biological process）。对意大利蝗转录组 unigene 进行 GO 分析发现，涉及分子功能的 unigenes 有31 392 条，其中分子功能（molecular function）、催化活性（catalytic activity）和结合（binding）的 unigenes 数量最多，分别是 4 455，2 762 和 2 553 条，其余均在 1 500条以下；涉及细胞组分的 unigenes 有 20 586 条，其中细胞组分（cellular component）、细胞膜（membrane）、细胞（cell）和细胞部分（cell part）的 unigenes 数量最多，分别是 2 909，1 657，1 502 和 1 502条，其余大部分在 1 500 条以下；与生物过程有关的 unigenes 有 57 014 条，其中生物过程（biological process）、代谢过程（metabolic process）、细胞过程（cellular process）、有机物代谢过程（organic substance metabolic process）的 unigenes 均在 2 000 条以上，分别为 3 629，2 483，2 381 和 2 022 条，其余均在 2 000 条以下（图 2）。

此外，所有 unigenes 经过 KOG 数据库功能预测和分类，共有 94 262 条基因被分为 26 类（图 3）。其中只有一般功能（general function prediction only）的 unigenes 数量最多，为 46 367 条（49.19%），是最大类群；其次分别是参与信号转导机制（signal transduction mechanisms）的 5 430 条（5.76%），参与细胞周期调控、细胞分裂和染色体分区（cell cycle control，cell division，chromosome partitioning）的 5 424条（5.75%），参与翻译后修饰、蛋白质转换、分子伴侣（posttranslational modification，protein turnover，chaperones）的 4 600 条（4.88%），参与染色质结构和动力学（chromatin structure and dynamics）的 3 909 条（4.15%），参与转录（transcription）的 2 569 条（2.73%），参与未知功能（function unknown）的 2 383 条（2.53%），参与糖类运输与代谢（carbohydrate transport and metabolism）的 2 275 条（2.41%），参与翻译、核糖体结构和生物合成（translation，ribosomal structure and biogenesis）的 2 202 条（2.34%），参与复制、重组和修复（replication，recombination and repair）的 2 151 条（2.28%），参与细胞内运输、分泌和膜泡运输（intracellular trafficking，secretion and vesicular transport）的 2 060 条（2.19%）。unigenes 数量在 1 000 ~ 2 000 之间的有氨基酸运输与代谢（amino acid transport and metabolism）、能量的产生和转换（energy production and conversion）、RNA 加工和修饰（RNA processing and modification）、脂类运输与代谢（lipid transport and metabolism）、次级代谢产物的生物合成、转运和分解（secondary metabolites biosynthesis，transport and catabolism）、细胞骨架（cytoskeleton）、无机离子转运和代谢（inorganic ion transport and metabolism）和细胞壁/膜/包膜的生物发生（cell wall/membrane/envelope biogenesis），分别为 1 917，1 838，1 808，1 764，1 711，1 583，1 131 和 1 049 条。其他的 unigenes 数量均少于 1 000 条，其中与核结构（nuclear structure）和细胞动力（cell motility）有关的 unigenes 更少，所占比例分别为 0.21% 和 0.07%。

使用 KEGG 注释系统进行 unigene 代谢途径分析，结果表明，23 666 条 unigenes 归属于 251 条通路，其中碳代谢（carbon metabolism）通路最多，为 468 条；含有 300 ~ 400 条 unigenes 的通路有：PI3K-Akt 信号途径（PI3K-Akt signaling pathway），367 条；氨基酸合成（biosynthesis of amino acids），349 条；含有 200 ~ 300 条 unigenes 的通路有：嘧啶代谢（pyrimidine metabolism），299 条；淀粉和蔗糖代谢（starch and sucrose metabolism），295 条；粘着斑（focal adhesion），281 条；吞噬体（phagosome）253 条；细胞周期（cell cycle），244 条；糖酵解途径（glycolysis /gluconeogenesis），243 条；RNA 运转（RNA transport），237 条；丙酮酸代谢（pyruvate metabolism），232 条；AMPK 信号途径（AMPK signaling pathway），226 条；内质网中的蛋白质合成（protein processing in endoplasmic reticulum），224 条；cAMP 信号途径（cAMP signaling pathway），219 条；cGMP-PKG 信号途径（cGMP-PKG signaling pathway），218 条；卵母细胞减数分裂（oocyte meiosis），216 条；MAPK 信号通道（MAPK signaling pathway），208 条。其他通路的 unigenes 数量均在 200 条以下，其中涉及性腺发育的通路有：胰岛素信

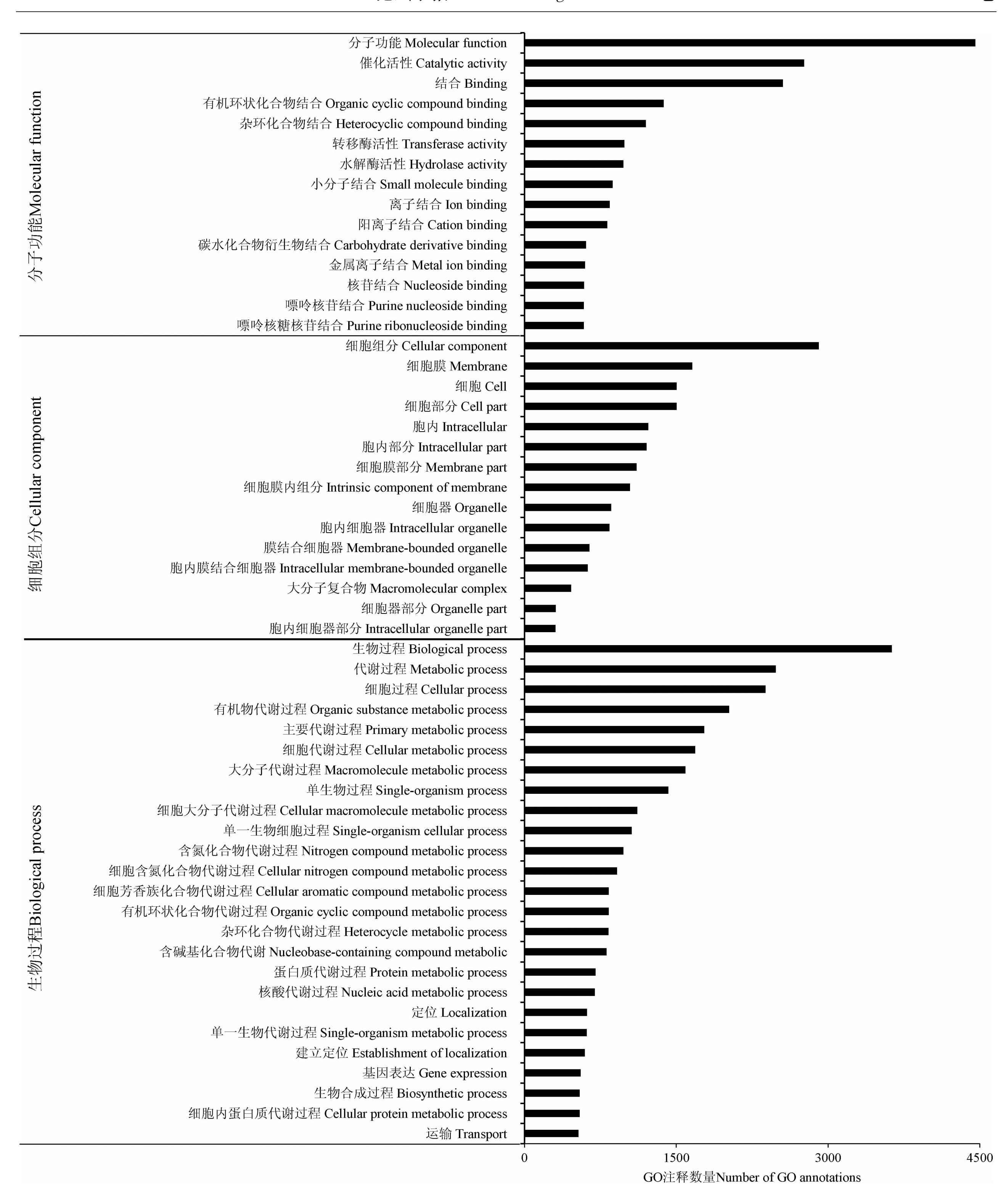


图 2 GO 分类

Fig. 2 Gene Ontology (GO) categories

号通路(insulin signaling pathway),185 条;Wnt 信号通路(Wnt signaling pathway),183 条;促性腺激素释放激素信号通路(GnRH signaling pathway),115 条;转化生长因子 β 信号通路(TGF-β signaling pathway),101 条;泛素介导的蛋白水解(ubiquitin mediated proteolysis),76 条;黄体酮介导的卵母细胞成熟(progesterone-mediated oocyte maturation),66 条;昆虫激素生物合成(insect hormone biosynthesis),5 条(表 4)。

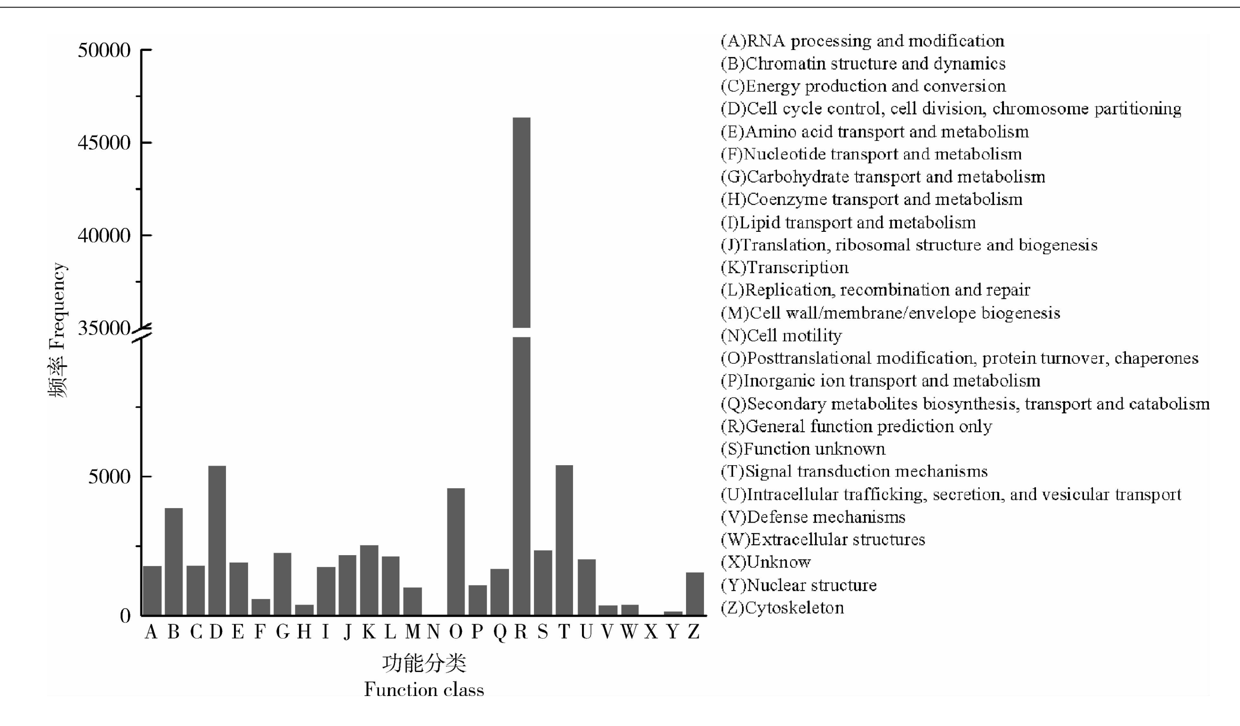


图 3　KOG 分类

Fig. 3　Karyotic Ortholog Groups（KOG）classification

表 4　意大利蝗转录组序列 KEGG 通路统计

Table 4　Statistics of KEGG pathways for *Calliptamus italicus* transcriptome

代谢通路 Pathways	通路 ID Pathway ID	基因数量 Number of genes
碳代谢 Carbon metabolism	ko01200	468
PI3K-Akt 信号途径 PI3K-Akt signaling pathway	ko04151	367
氨基酸合成 Biosynthesis of amino acids	ko01230	349
嘧啶代谢 Pyrimidine metabolism	ko00240	299
淀粉和蔗糖代谢 Starch and sucrose metabolism	ko00500	295
粘着斑 Focal adhesion	ko04510	281
吞噬体 Phagosome	ko04145	253
细胞周期 Cell cycle	ko04110	244
糖酵解途径 Glycolysis/Gluconeogenesis	ko00010	243
RNA 运转 RNA transport	ko03013	237
丙酮酸代谢 Pyruvate metabolism	ko00620	232
AMPK 信号途径 AMPK signaling pathway	ko04152	226
内质网中的蛋白质合成 Protein processing in endoplasmic reticulum	ko04141	224
cAMP 信号途径 cAMP signaling pathway	ko04024	219
cGMP-PKG 信号途径 cGMP-PKG signaling pathway	ko04022	218
卵母细胞减数分裂 Oocyte meiosis	ko04114	216
MAPK 信号通道 MAPK signaling pathway	ko04010	208
其他通路 Others	–	<200
胰岛素信号通路 Insulin signaling pathway	ko04910	185
Wnt 信号通路 Wnt signaling pathway	ko04310	183
促性腺激素释放激素信号通路 GnRH signaling pathway	ko04912	115
转化生长因子 β 信号通路 TGF-β signaling pathway	ko04350	101
泛素介导的蛋白水解 Ubiquitin mediated proteolysis	ko04120	76
黄体酮介导的卵母细胞成熟 Progesterone-mediated oocyte maturation	ko04914	66
昆虫激素生物合成 Insect hormone biosynthesis	ko00981	5
通路总数 Total	–	23 666

2.4 意大利蝗性腺发育相关基因的注释

对意大利蝗雌雄成虫基因表达量进行分析，筛选得到24 653 个差异表达基因，其中12 954 个基因在雄虫表达上调，11 699 个基因在雌虫表达上调，且10 824 个基因仅在雌虫表达。基于差异表达基因的功能注释，结合已有的相关文献报道，在转录组数据中筛选到涉及性腺发育过程的重要功能基因（表5）。

表5 意大利蝗转录组中与性腺发育相关的基因

Table 5 Gonadal development related genes in *Calliptamus italicus* transcriptome

基因 Genes	物种 Matched species	长度(bp) Length	*E* 值 *E* value	一致性(%) Identity
性腺发育 Gonadal development				
vg	大田鳖 *Lethocerus deyrollei*	4 967	0	31.199
vgr	美洲大蠊 *Periplaneta americana*	5 451	0	45.163
sox 21b	黑脉金斑蝶 *Danaus plexippus*	350	6.27E-24	97.778
PREDICTED: *testis-specific serine/threonine-protein kinase 2*	松叶蜂 *Neodiprion lecontei*	4 035	0	84.848
PREDICTED: *testis-specific serine/threonine-protein kinase 4-like*	松叶蜂 *Neodiprion lecontei*	1 501	2.08E-150	68.667
Spermatogenesis-associated protein 5	内华达古白蚁 *Zootermopsis nevadensis*	2 703	0	80.658
Spermatogenesis-associated protein 22	内华达古白蚁 *Zootermopsis nevadensis*	2 289	3.77E-29	44.526
Vasa-like gene	沙漠蝗 *Schistocerca gregaria*	3 792	0	90.167
昆虫激素生物合成 Insect hormone biosynthesis				
Dehydrogenase/reductase SDR family member 11	内华达古白蚁 *Zootermopsis nevadensis*	874	2.36E-68	52.792
Juvenile hormone acid methyltransferase	沙漠蝗 *Schistocerca gregaria*	2 210	3.22E-54	50.296
Juvenile hormone epoxide hydrolase 1	内华达古白蚁 *Zootermopsis nevadensis*	1 025	1.26E-154	69.072
Juvenile hormone esterase	黄星桑天牛 *Psacothea hilaris*	2 406	9.84E-88	33.789
PREDICTED: *cytochrome P450 307a1*	赤拟谷盗 *Tribolium castaneum*	4 345	2.93E-109	67.557
黄体酮介导的卵母细胞成熟 Progesterone-mediated oocyte maturation				
Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase	内华达古白蚁 *Zootermopsis nevadensis*	2 402	1.86E-77	48.084
Mitotic checkpoint serine/threonine-protein kinase BUB1	内华达古白蚁 *Zootermopsis nevadensis*	3 883	6.96E-77	45.912
Phosphatidylinositol-4, 5-bisphosphate 3-kinase catalytic subunit delta isoform	内华达古白蚁 *Zootermopsis nevadensis*	4 606	0	80.1678
PREDICTED: *cytoplasmic polyadenylation element-binding protein 2 isoform X4*	丽蝇蛹集金小蜂 *Nasonia vitripennis*	2 073	1.14E-09	100
PREDICTED: LOW QUALITY PROTEIN: *adenylate cyclase type 9-like*	大蜜蜂 *Apis dorsata*	5 733	0	57.837
促性腺激素释放激素信号通路 GnRH signaling pathway				
Adenylate cyclase type 2	短管赤眼蜂 *Trichogramma pretiosum*	4 372	3.09E-174	80.958
Calmodulin	柑橘凤蝶 *Papilio xuthus*	2 603	2.97E-95	98.675
Early growth response 1	东亚飞蝗 *Locusta migratoria*	3 990	0	98.288
Epidermal growth factor receptor	内华达古白蚁 *Zootermopsis nevadensis*	7 265	0	66.482
Mitogen-activated protein kinase kinase kinase 1	内华达古白蚁 *Zootermopsis nevadensis*	4 628	0	41.745
PREDICTED: *gonadotropin-releasing hormone* Ⅱ *receptor-like*	芜菁叶蜂 *Athalia rosae*	1 016	8.13E-21	75.385
PREDICTED: LOW QUALITY PROTEIN: *adenylate cyclase type 9-like*	大蜜蜂 *Apis dorsata*	5 733	0	57.837
Transcription factor AP-1	内华达古白蚁 *Zootermopsis nevadensis*	2 104	1.13E-35	69.792
Tyrosine-protein kinase Src64B	内华达古白蚁 *Zootermopsis nevadensis*	2 558	0	81.887

续表 5 Table 5 continued

基因 Genes	物种 Matched species	长度(bp) Length	*E* 值 *E* value	一致性(%) Identity
卵母细胞减数分裂 Oocyte meiosis				
Aurora kinase A	袋獾 *Sarcophilus harrisii*	4 575	1.85E－27	72.824
Calmodulin	柑橘凤蝶 *Papilio xuthus*	2 603	2.97E－95	98.675
Cell division cycle protein 20-like protein	内华达古白蚁 *Zootermopsis nevadensis*	2 765	5.69E－71	34.459
G1/S-specific cyclin-E	内华达古白蚁 *Zootermopsis nevadensis*	3 106	1.18E－132	53.102
Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase	内华达古白蚁 *Zootermopsis nevadensis*	2 402	1.86E－77	48.084
Mitotic checkpoint serine/threonine-protein kinase BUB1	内华达古白蚁 *Zootermopsis nevadensis*	3 883	6.96E－77	45.912
PREDICTED：*cytoplasmic polyadenylation element-binding protein 2 isoform X4*	丽蝇蛹集金小蜂 *Nasonia vitripennis*	2 073	1.14E－09	100
PREDICTED：LOW QUALITY PROTEIN：*adenylate cyclase type 9-like*	大蜜蜂 *Apis dorsata*	5 733	0	57.837
Protein phosphatase 2A regulatory subunit A, putative	点蜂缘蝽 *Riptortus pedestris*	282	1.65E－40	75.269
Serine/threonine-protein phosphatase PP1	内华达古白蚁 *Zootermopsis nevadensis*	4 443	4.67E－178	89.007
细胞周期 Cell cycle				
CDK5 regulatory subunit-associated protein 1	芜菁叶蜂 *Athalia rosae*	2 356	0	66.160
Cell division cycle 2-like protein kinase	内华达古白蚁 *Zootermopsis nevadensis*	1 735	0	60.405
Cell division cycle 5-like protein	灰翅胡椒鲷 *Cephus cinctus*	5 244	0	76.903
Cyclin dependent kinase 1	红火蚁 *Solenopsis invicta*	6 048	2.48E－17	82.645
Cyclin dependent kinase 2	斑翅果蝇 *Drosophila suzukii*	2 388	6.64E－08	77.143
MAPK 信号通路 MAPK signal pathway				
14-3-3 zeta	拟稻蝗 *Gesonula punctifrons*	4 122	0	83.967
14-3-3 protein epsilon	沙漠蝗 *Schistocerca gregaria*	6 158	3.54E－167	94.465
Mitogen-activated protein kinase 1-like	美洲鲎 *Limulus polyphemus*	2 043	4.31E－120	74.901
Mitogen-activated protein kinase kinase kinase 15	印度跳蚁 *Harpegnathos saltator*	5 571	0	75.481
泛素介导的蛋白水解 Ubiquitin mediated proteolysis				
E3 ubiquitin-protein ligase SIAH1	斑点雀鳝 *Lepisosteus oculatus*	6 831	1.36E－169	81.744
Polyubiquitin-C-like	橄榄果蝇 *Bactrocera oleae*	260	3.40E－13	88.571
Ubiquitin carboxyl-terminal hydrolase isozyme L5	美洲东部熊蜂 *Bombus impatiens*	4 700	2.51E－11	77.551

3　讨论

对于缺少基因组信息的非模式生物，通过转录组测序，可以从特定组织中获得参与特定生物学过程的功能基因信息（Künstner *et al*., 2010；Jung *et al*., 2011；Teaniniuraitemoana *et al*., 2014）。本研究选取性成熟期（交配前期）意大利蝗性腺组织进行转录组测序和分析，以期获得意大利蝗生殖相关基因信息，共获得 718 872 条高质量 unigenes，其中有 254 597 条 unigenes 在 Nr，KOG，KEGG，Swiss-Prot 和 GO 等数据库中被成功注释。KEGG 分析表明，23 666 条 unigenes 归属于 251 条代谢通路，包括一些在性腺发育和成熟过程中必要的信号通路（Meng *et al*., 2015），如与内分泌系统（endocrine system）相关的胰岛素信号通路、促性腺激素释放激素信号通路、黄体酮介导的卵母细胞成熟途径，与信号转导（signal transduction）相关的 PI3K-Akt 信号途径、AMPK 信号途径、cAMP 信号途径、cGMP-PKG 信号途径、MAPK 信号通道、Wnt 信号通路、转化生长因

子β信号通路等。同时,亦鉴定到参与信号转导的相关基因,如腺苷酸环化酶2(*adenylate cyclase type 2*)、钙调素(*calmodulin*)、磷脂酰肌醇-4,5-二磷酸激酶催化亚基(*phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta isoform*)、酪氨酸蛋白激酶src64b(*tyrosine-protein kinase Src64B*)基因等,这些基因在意大利蝗性腺组织中表现出较高的表达量。相关代谢通路和基因的鉴定表明,内分泌调节系统和信号转导系统在意大利蝗性腺发育和功能调节中具有重要作用。通过生物信息学分析,有助于从大规模转录组数据库中识别潜在的功能基因,为意大利蝗生殖调控机制研究提供丰富的基因信息资源。

本研究筛选得到24 653个雌雄差异表达基因,这些基因为研究意大利蝗生殖调控机制提供了重要的基因信息。卵巢发育和卵子发生是决定雌虫繁殖力的关键因素,是种群繁衍的重要基础。卵黄发生是卵子发生中的关键时期,是昆虫生殖调控的核心,卵母细胞中卵黄蛋白的迅速积累,可促进卵母细胞和卵巢的发育成熟,亦可为胚胎发育提供营养物质,有利于提高孵化率及幼虫生长发育,对昆虫种群数量有显著影响(Muhammad *et al.*, 2014)。卵黄发生受多种因素的调控,其中卵黄蛋白(*vitellogenin*, *Vg*)及其受体(*vitellogenin receptor*, *VgR*)基因的表达是卵黄发生的基础,对昆虫卵巢成熟起着至关重要的作用,也是研究控制害虫的潜在靶标(Muhammad *et al.*, 2014; Roy-Zokan *et al.*, 2015; Upadhyay *et al.*, 2016)。本研究中,Vg和VgR在意大利蝗卵巢转录组中表达上调,揭示其在卵巢发育中的重要作用。在意大利蝗转录组中亦发现大量与精子发生相关的基因,如*sox 21b*、睾丸特异性丝氨酸/苏氨酸蛋白激酶(*testis-specific serine/threonine-protein kinases*)、精子发生相关蛋白(*spermatogenesis-associated protein*)基因等,这些基因在精巢组织中高表达,提示其参与了精巢分化和精子发生的调控。*vasa*基因是DEAD-box家族基因的一员,广泛存在于所有真核生物和大多数原核生物中,是生殖系统发育的重要调控因子(Cordin *et al.*, 2006)。*vasa*基因突变将导致生殖细胞分化和卵细胞发育的异常,影响正常的卵子发生进程,造成雌性个体不育(Styhler *et al.*, 1998)。研究表明,*vasa*仅在沙漠蝗*Schistocerca gregaria*精巢和卵巢中特异表达,*vasa* mRNA在早期卵子发生中高表达,而在发育成熟的卵中不表达(Chang *et al.*, 2002)。本研究在意大利蝗精巢、卵巢转录组中均发现*vasa*基因,且表达量较高,揭示*vasa*基因在意大利蝗性腺发育中发挥重要作用。此外,亦获得参与细胞周期、泛素介导的蛋白水解等通路的相关基因,研究表明,这些信号通路均参与一系列生殖调控活动,包括配子发生、性腺发育等(高洁, 2013; Meng *et al.*, 2015; Peng *et al.*, 2015; Yang *et al.*, 2016)。

蝗虫性腺发育还受到激素等物质的调控,其中,保幼激素(juvenile hormone, JH)对于卵黄发生及卵母细胞成熟具有重要的调节作用(Guo *et al.*, 2014)。本研究在意大利蝗性腺转录组中注释得到参与保幼激素合成和代谢的功能基因,如保幼激素酸甲基转移酶(*juvenile hormone acid methyltransferase*)、保幼激素环氧化物水解酶1(*juvenile hormone epoxide hydrolase 1*)、保幼激素酯酶(*juvenile hormone esterase*)基因等,这些基因的鉴定和功能分析将在分子水平上促进意大利蝗生殖内分泌研究。

本研究利用多种数据库对所获得的unigene进行BLASTX比对分析,成功注释254 597条unigene,并筛选得到参与意大利蝗性腺发育和配子发生的相关功能基因,这些基因在意大利蝗生殖调控中的确切作用还有待于进一步研究。然而,由于unigenes整体长度较短、昆虫基因信息仍相对缺乏等原因,仍有64.58%(464 275条unigene)的unigene没有获得注释信息,其中可能包含新的与性腺发育和生殖调控相关的功能基因,这些基因的鉴定和功能分析还有待于进一步研究。

参考文献(References)

Chang CC, Dearden P, Akam M, 2002. Germ line development in the grasshopper *Schistocerca gregaria*: *vasa* as a marker. *Dev. Biol.*, 252(1): 100-118.

Cordin O, Banroques J, Tanner NK, Linder P, 2006. The DEAD-box protein family of RNA helicases. *Gene*, 367: 17-37.

Gao J, 2013. Comparative Transcriptome Analysis for Gonad of the *Scylla paramamosain*. MSc Thesis, Jimei University, Xiamen, Fujian. [高洁, 2013. 拟穴青蟹性腺差异的转录组学分析. 福建厦门: 集美大学硕士学位论文]

Guo W, Wu ZX, Song JS, Jiang F, Wang ZM, Deng S, Walker VK, Zhou ST, 2014. Juvenile hormone-receptor complex acts on *Mcm4* and *Mcm7* to promote polyploidy and vitellogenesis in the migratory locust. *PLoS Genet.*, 10(10): e1004702.

Hunter DM, Latchininsky AV, Abashidze E, Gapparov FA, Nurzhanov AA, Medetov MZ, Tufliev NX, 2016. The efficacy of *Metarhizium acridum* against nymphs of the Italian locust, *Calliptamus italicus* (L.) (Orthoptera: Acrididae) in Uzbekistan and Georgia. *J.*

Orthop. Res., 25(2): 61 - 65.

Jin YL, Cong B, Wang LY, Zhang HY, Dong H, 2015. An analysis of the transcriptome of *Epacromius coerulipes* (Orthoptera: Acrididae). *Acta Entomol. Sin.*, 58(8): 817 - 825. [金永玲, 丛斌, 王丽艳, 张海燕, 董辉, 2015. 大垫尖翅蝗转录组分析. 昆虫学报, 58(8): 817 - 825]

Jung H, Lyons RE, Dinh H, Hurwood DA, McWilliam S, Mather PB, 2011. Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): *de novo* assembly, annotation and marker discovery. *PLoS ONE*, 6(12): e27938.

Kim IW, Markkandan K, Lee JH, Subramaniyam S, Yoo S, Park J, Hwang JS, 2016. Transcriptome profiling and in silico analysis of the antimicrobial peptides of the grasshopper *Oxya chinensis sinuosa*. *J. Microbiol. Biotechnol.*, 26(11): 1863 - 1870.

Künstner A, Wolf JBW, Backström N, Whitney O, Balakrishnan CN, Day L, Edwards SV, Janes DE, Schlinger BA, Wilson RK, Jarvis ED, Warren WC, Ellegren H, 2010. Comparative genomics based on massive parallel transcriptome sequencing reveals patterns of substitution and selection across 10 bird species. *Mol. Ecol.*, 19 (Suppl. 1): 266 - 276.

Meng XL, Liu P, Jia FL, Li J, Gao BQ, 2015. *De novo* transcriptome analysis of *Portunus trituberculatus* ovary and testis by RNA-Seq: identification of genes involved in gonadal development. *PLoS ONE*, 10(6): e0128659.

Muhammad T, Yoshiki N, Azza ME, Makio T, 2014. Regulation of vitellogenin genes in insects. *Entomol. Sci.*, 17: 269 - 282.

Oppenheim SJ, Baker RH, Simon S, DeSalle R, 2015. We can't all be supermodels: the value of comparative transcriptomics to the study of non-model insects. *Insect Mol. Biol.*, 24(2): 139 - 154.

Peng JX, Wei PY, Zhang B, Zhao YZ, Zeng DG, Chen XL, Li M, Chen XH, 2015. Gonadal transcriptomic analysis and differentially expressed genes in the testis and ovary of the Pacific white shrimp (*Litopenaeus vannamei*). *BMC Genom.*, 16(1): 1006.

Popova EN, Semenov SM, Popov IO, 2016. Assessment of possible expansion of the climatic range of Italian locust (*Calliptamus italicus* L.) in Russia in the 21st century at simulated climate changes. *Russ. Meteorol. Hydrol.*, 41(3): 213 - 217.

Qiu ZY, Liu F, Lu HM, Huang Y, 2017. Characterization and analysis of a *de novo* transcriptome from the pygmy grasshopper *Tetrix japonica*. *Mol. Ecol. Resour.*, 17(3): 381 - 392.

Qiu ZY, Liu F, Lu HM, Yuan H, Zhang Q, Huang Y, 2016. *De novo* assembly and characterization of the transcriptome of grasshopper *Shirakiacris shirakii*. *Int. J. Mol. Sci.*, 17: 1110.

Ren JL, Tu XB, Ge J, Zhao L, Zhang ZH, 2016. Influence of temperature on the development, reproduction, and life table of *Calliptamus italicus* (L.) (Orthoptera: Acridoidea). *J. Asia-Pac. Entomol.*, 19(1): 203 - 207.

Roy-Zokan EM, Cunningham CB, Hebb LE, McKinney EC, Moore AJ, 2015. Vitellogenin and vitellogenin receptor gene expression is associated with male and female parenting in a subsocial insect. *Proc. R. Soc. B*, 282: 20150787.

Styhler S, Nakamura A, Swan A, Suter B, Lasko P, 1998. *Vasa* is required for GURKEN accumulation in the oocyte, and is involved in oocyte differentiation and germline cyst development. *Development*, 125(9): 1569 - 1578.

Teaniniuraitemoana V, Huvet A, Levy P, Klopp C, Lhuillier E, Gaertner-Mazouni N, Gueguen Y, Moullac GL, 2014. Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: identification of potential sex differentiation and sex determining genes. *BMC Genom.*, 15(1): 491.

Upadhyay SK, Singh H, Dixit S, Mendu V, Verma PC, 2016. Molecular characterization of vitellogenin and vitellogenin receptor of *Bemisia tabaci*. *PLoS ONE*, 11(5): 1 - 18.

Wang DM, Li S, Zhang YJ, Jashenko R, Ji R, 2016. Response of the discontinuous gas exchange cycle (DGC) duration in *Calliptamus italicus* (Orthoptera: Acrididae) to high temperature stress. *Acta Entomol. Sin.*, 59(5): 516 - 522. [王冬梅, 李爽, 张永军, Jashenko R, 季荣, 2016. 意大利蝗不连续气体交换循环(DGC)呼吸周期历时对高温胁迫的响应. 昆虫学报, 59(5): 516 - 522]

Wang H, He XQ, Ji R, 2010. Selection mechanisms of *Calliptamus italicus* on four different host plants. *Chin. J. Ecol.*, 29(12): 2401 - 2407. [王晗, 何雪青, 季荣, 2010. 意大利蝗对四种寄主植物的选择机制. 生态学杂志, 29(12): 2401 - 2407]

Xiang M, Fan TS, Hu HX, Yu F, Ji R, Wang H, 2017. Effects of short-term exposure to high temperature on the survival and fecundity of *Calliptamus italicus* (Orthopera: Acrididae). *Chin. J. Appl. Entomol.*, 54(3): 426 - 433. [向敏, 樊泰山, 扈鸿霞, 于非, 季荣, 王晗, 2017. 短时高温对意大利蝗存活和生殖的影响. 应用昆虫学报, 54(3): 426 - 433]

Xue ZP, Zhang Q, Yasen S, Wang GJ, Abudouwaili Y, Xiao HW, 2010. Food selection of *Calliptamus italicus* and loss assessment for the grassland. *Plant Prot.*, 36(1): 95 - 98. [薛智平, 张泉, 牙森・沙力, 王广君, 阿不都外力・伊玛木, 肖宏伟, 2010. 意大利蝗取食特性及损失估计研究. 植物保护, 36(1): 95 - 98]

Yang D, Yin C, Chang YQ, Dou Y, Hao ZL, Ding J, 2016. Transcriptome analysis of male and female mature gonads of Japanese scallop *Patinopecten yessonsis*. *Genes Genom.*, 38(11): 1041 - 1052.

Yang J, 2013. Comparative Transcriptome of *Atractomorpha sinensis* at Three Developmental Stages and Mitochondrial Transcript Mapping. MSc Thesis, Shaanxi Normal University, Xi'an. [杨婧, 2013. 短额负蝗三种虫态的比较转录组及线粒体转录组作图研究. 西安: 陕西师范大学硕士学位论文]

Zhang B, Zhang W, Nie RE, Li WZ, Segraves KA, Yang XK, Xue HJ, 2016. Comparative transcriptome analysis of chemosensory genes in two sister leaf beetles provides insights into chemosensory speciation. *Insect Biochem. Mol. Biol.*, 79: 108 - 118.

Zhang S, 2015. Identification and Expression Profile Analysis of Odorant-binding Proteins Genes Based on Transcriptome in Grasshopper, *Oedaleus asiaticus*. MSc Thesis, Inner Mongolia Agricultural University, Hohhot. [张硕, 2015. 基于转录组的亚洲小车蝗气味结合蛋白基因的鉴定及其表达谱分析. 呼和浩特: 内蒙古农业大学硕士学位论文]

Zhang XH, Li YL, Zhang KJ, Hou J, Yin H, 2016. Cloning and expression analysis of the hexamerin subunit type 2 (Hex2) gene from the grasshopper *Calliptamus italicus* (Orthoptera: Catantopidae). *Acta Entomol. Sin*., 59(2): 156-163. [张晓红, 李艳丽, 张凯建, 侯娟, 印红, 2016. 意大利蝗存储蛋白 Hex2 基因的克隆与表达分析. 昆虫学报, 59(2): 156-163]

Zhao Z, Hao XL, Li N, Ren BZ, 2008. The dynamic development of spermary and ovary of six species of dominant locusts in Siping area of Jilin Province, China. *J. Jilin Agric. Univ*., 30(5): 677-681. [赵卓, 郝锡联, 李娜, 任炳忠, 2008. 四平地区 6 种蝗虫精巢和卵巢发育动态. 吉林农业大学学报, 30(5): 677-681]

Zhao Z, Xi GS, 2005. The seasonal dynamics of ovarian development of six dominant species of locusts in Siping area of Jilin. *Chin. Bull. Entomol*., 42(5): 524-527. [赵卓, 奚耕思, 2005. 吉林四平地区六个蝗虫优势种卵巢发育的季节动态. 昆虫知识, 42(5): 524-527]

Zhao ZW, Zhang YC, Cao GC, Zhang ZH, 2013. Influence of temperature on the development of *Calliptamus italicus*. *Chin. J. Appl. Entomol*., 50(2): 466-473. [赵忠伟, 张英财, 曹广春, 张泽华, 2013. 温度对意大利蝗生长发育的影响. 应用昆虫学报, 50(2): 466-473]

(责任编辑: 马丽萍)